CENTRE FOR
SYSTEMS
SCIENCE

LCCR

Simon Fraser University, Burnaby, B.C., Canada V5A 1S6 / (604) 291-4704

Natural Metamers

Mark S. Drew and Brian V. Funt

LCCR TR 90-05

February 1991

92-05791/

Natural Metamers

Mark S. Drew and Brian V. Funt
School of Computing Science
Simon Fraser University
Vancouver, B.C.
Canada V5A 1S6
(604) 291-4682

©1990 M.S. Drew and B.V. Funt

Abstract

Several methods are proposed and tested for recovering a complete spectrum given only a color camera's RGB measurement of it as input. The recovered spectrum can be used to generate tristimulus values, so in essence, we provide a new method of converting color camera signals to tristimulus coordinates. The importance of this result lies in the fact that for accurate color reproduction, the colors reproduced must be metamers of the originals; and two spectra having equal tristimulus values will be metameric while in general two spectra causing equal camera responses will not be metameric. We compare the accuracy of our method of converting to tristimulus coordinates from RGB signals to one based on Horn's work on the exact reproduction of colored images. The new method relies on expanding the color signal as a series in terms of a linear combination of basis functions. The results show that the best basis comes from performing a principal component analysis in color signal space, since such basis functions recover the most "natural" color signals that are metameric with the camera inputs.

1 Introduction

Two light spectra that *look* the same even though the spectra themselves are different are called metamers and are said to be metameric to one another. Metamers occur because the eye, or a camera, collapses all the spectral information into a set of just three numbers, the RGB signal in a camera or presumably a similar set of three excitations in a normal human with trichromatic vision. Metamers cause a problem whenever one shifts from one system, the human one say, to another system such as a camera system, because what is a metamer for one in general will not be for the other.

As an example, consider three different objects with artificially constructed spectra (given in [24], p. 171) that would all look the same to a human when viewed under a standard daylight—pale tangerine. One quantifies this intuitive idea by assigning tristimulus values X, Y, Z (or simply XYZ) to a color. These values are related to the proportion of color primaries that must be added together to make the color. From XYZ one defines the chromaticity of the color by the pair (x, y) where $x = X/(X + Y + Z)$ and $y = Y/(X + Y + Z)$. Chromaticities are usually displayed on a diagram devised by the CIE in 1931. For our tangerine color the tristimulus values are $(129.06, 100.00, 44.20)$ and the chromaticity is $(x, y) = (0.4691, 0.3643)$.

The same objects under the same daylight illumination, to a color camera, all appear different from one another, however. Specifically, the values of the RGB signal produced by the camera are $(1.490, 0.901, 0.388)$ for the first, $(1.476, 0.838, 0.477)$ for the second, and $(1.189, 1.150, 0.422)$ for the last, where an equi-energy white spectrum produces an RGB signal of $(1, 1, 1)$. If these camera signals are now reproduced on a terminal or in a color print the chromaticities for each object will be shifted, possibly by quite a lot depending on the color correction system in use. On one system we tried, passing these RGBs [1] directly to the output device resulted in the colors tangerine, salmon, and a pale green tan. The camera sees the spectra as discernibly different, even though the eye does not.

Turning the situation around, suppose that the camera sees three different spectra as metameric. I.e., their RGB signals are the same, as far as the camera system is concerned. What XYZ value are we to assign to the RGB signal? In this paper, we seek to answer this question. Of course, any answer must be a compromise because there should be three different XYZ triples, not just one.

The problem of metamerism has a long history in color science, although it is not always recognized in computer graphics and computer vision as part of the problem in accurate

[1] Throughout this paper we denote by 'RGB' the camera system signal (not three tristimulus values, as in [24]).

color reproduction. A visible-light signal is characterized by the spectral power distribution (SPD) of relative intensities of the signal measured over many (usually evenly-spaced) wavelength intervals. The SPD can be thought of as an approximation of an analytic function of wavelength or, if the spectrum is sampled at N wavelength intervals, as a vector in an N-dimensional space. The signal that arrives at the lens (either camera or human) is composed of the light spectrum illuminating a surface multiplied by the surface spectral reflectance function. The resulting signal is variously called the *color signal* in computer vision (see, e.g., [23]) or the *object-color stimulus* in color science [24].

When filtered by color filters and viewed by camera (or eye) sensors, the color signal for each pixel is condensed into three RGB values that carry the intensity and color information for the image. (We shall sidestep the issue of just how RGB information is created in the eye by relying in the main on psychophysical color-matching functions [24] to characterize human sensors.) The problem of metamerism arises because the mapping from SPD to RGB is many-to-one: there are usually a great many possible color signals that could have resulted in exactly the same set of sensor responses.

In terms of color reproduction this would create no problem if camera sensors were fooled by the same metamers as the human eye, i.e. if *camera metamers* were the same as *eye metamers* (*cf.* [15]). However, as Horn has pointed out [9], this is only possible when the sensor curves are a homogeneous linear transformation away from human sensor response curves (or, more practically, from human color-matching functions). But this linear relationship between response curves for machine and human vision does not in general hold.

Horn [9] gives a straightforward method for determining the linear transformation of the color-matching functions that "best" corresponds to the actual sensor response curves. His prescription implies, as well, a simple linear transformation of the camera RGB signal into corresponding tristimulus values XYZ. (Here, we call this linear transformation '**Method A**'.) It is the tristimulus values that form the input to the chain of transformations leading to color reproductions, either video or printed [9, 21]. Horn's approach is an attempt to be algorithmic rather than rely in large measure on practitioners in the color reproduction field to apply their experience and rules of thumb to the color correction problem.

Our interest in this topic stems from our work on the computational vision problem of separating the color signal into its component factors, illumination and reflectance [8]. Using the statistical properties of typical daylights and naturally occurring reflectance functions, we showed that it is indeed possible to disentangle the illumination from the reflectance, provided that one has available the entire SPD color signal function, not just an RGB signal. This information turns out to be available from lens systems with nonvanishing chromatic aberration [7], but is generally not available from standard video systems. For our purposes, then, it is necessary somehow to *produce* an entire color signal from just the sensor RGB

inputs, as a front end to our color signal separation algorithm.

Here we run into just the metamerism problem addressed in the color science field: which of the many available metamers, by definition all having the same sensor RGB, is the actual spectrum producing that RGB?

In our work on the color-signal-separation algorithm [8], we used finite dimensional linear models to reduce the amount of information necessary to characterize an illumination spectrum or a surface spectral reflectance function by capturing all the necessary information in a few weights of a small number of terms in a linear expansion in terms of a set of basis vectors. Here we use a similar idea by writing the color signal as a weighted sum of basis vectors, and then truncating the series as a means of producing a set of equations that can be solved for the basis function weights given the sensor RGB.

By generating a complete color-signal spectrum, we automatically also produce tristimulus values XYZ; and in fact the scheme is then similar to Method A in that the XYZ values are a simple linear transformation away from the input sensor RGB values. However, it turns out that for a good choice of basis vectors (we examine three different alternatives) our scheme (called **Method B**) produces XYZ values that are much closer to those of the original color signal. In fact, while Method A works reasonably well for sensor response curves that are nearly a linear transformation away from the color-matching functions (although only 40% as well as our method), it does not work at all well for real, bandpass filters. The accuracy of Method B, however, remains undiminished in these cases.

As well, since we are producing XYZ values by way of generating a complete color signal SPD, we can examine the SPD produced for appropriate physical characteristics. In particular, we find that it is simplest to throw away any spectra that have any negative components. This turns out to be 2% of the pixels examined, and can be accommodated easily by interpolation. The intermediate step of generating a complete SPD thus amounts to a useful tool for discarding those XYZs generated that would result from impossible spectra. This validation step is not available in Method A.

What we attack here is simply the first stage in the rather lengthy procedure for producing reproductions of color images, *viz.* the input conversion process for converting from sensor RGB values to tristimulus values XYZ (*cf.* [9, 21]). But on the way, we show that it is possible to recover the complete color-signal spectrum from the sensor values alone, with reasonable accuracy. We accomplish this by developing a set of basis functions for color signals by considering their statistical properties as vectors in an N-dimensional space.

Testing the method with 2052 synthetic, typical color signals, we found a median error for the color signals arrived at from just the sensor response RGB, as compared with the actual color signals, of 4.5 CIELUV units in 3-dimensional uniform color space.

2 Metamerism and Tristimulus Values

In reference [9], Horn is concerned with ascertaining the conditions for each step of the process of color image reproduction that guarantee that the resulting colors cannot be distinguished by humans, as well as with developing some indicators of the accuracy of reproduction when the exactness conditions are not met. Here we address just the first step of the reproduction process, that of converting the input sensor RGB responses into a best version of tristimulus values XYZ. The measure of accuracy we shall use is the CIELUV unit in uniform color space (see [24, p. 828]); this unit accounts for both the chromaticity difference between spectra and the luminance difference in absolute intensity.

If we generate a $C(\lambda)$ from the sensor RGB that matches the RGB when filtered by the camera filters, then we have generated a *metamer* with respect to the color filters that produced the RGB. Our problem consists in regenerating the one metamer among all the choices available that best matches (in CIELUV space) the actual color signal that produced the RGB. All the candidate metamers will have the same RGB (i.e., they will be camera metamers), but in general each will have a different XYZ (i.e. they will not be eye metamers).

Suppose that the camera has color filters $R_k(\lambda)$, so that the RGB sensor responses, denoted ρ_k, are given by

$$\rho_k = \int C(\lambda)\, R_k(\lambda)\, d\lambda \;,\; k = 1 \ldots 3 \;, \tag{1}$$

employing for convenience the function notation rather than vector notation. Then any $C(\lambda)$ spectrum is a camera metamer provided it has the same values for ρ_k. [2].

Tristimulus values XYZ are determined in an analogous manner, but we use the 1931 CIE standard observer color-matching functions instead of the sensor response curves R_k. These curves are commonly called $\bar{x}$, $\bar{y}$, $\bar{z}$, but for convenience we shall denote the collection of three curves by $\bar{x}_k$ and the collection X,Y,Z by X_k:

$$X_k = \int C(\lambda)\, \bar{x}_k(\lambda)\, d\lambda \;. \tag{2}$$

Our task is to generate from the ρ_k those values X_k that best match the X_k that would have been produced from the actual color signal. For computational vision applications, it would also be of use to know just what $C(\lambda)$ produced the measured ρ_k. We look at the second issue in the next section; here we set out a method for developing a set of X_k from the ρ_k that is based on Horn's work.

[2]Note that to determine the correct R_k it may be necessary to calibrate the camera system carefully for contributions from the lens system, the digitizer, etc. (*cf.* [15])

Reference [9] is concerned with determining strictures on the image reproduction system for the exact reproduction of colors. Horn shows that if the image sensor curves R_k are a linear transform of the color-matching functions $\bar{x}_k$, then camera metamers are the same as eye metamers. As well, he develops the linear transform of the $\bar{x}_k$ that "best" fits the sensor curves R_k. To do so, he assumes that the actual response curves R_k can be approximated by a linear transformation of the $\bar{x}_k$:

$$R_i(\lambda) \simeq \sum_{k=1}^{3} a_{ik}\, \bar{x}_k(\lambda)\,. \tag{3}$$

Then by minimizing the squared differences between the actual response curves R_k and the linear transform curves, the weights a_{ik} can be solved for:

$$\sum_{k=1}^{3} a_{ik}\, q_{kj} \;=\; v_{ij}\,, \tag{4}$$

where

$$q_{kj} \;=\; \int x_k(\lambda)\, \bar{x}_j(\lambda)\, d\lambda$$

$$v_{ij} \;=\; \int R_i(\lambda)\, x_j(\lambda)\, d\lambda$$

or, in matrix notation,

$$A \;=\; V\, Q^{-1} \tag{5}$$

assuming a nonsingular matrix Q. The approximation eqn. (3) amounts to a projection of the sensor curves onto the space spanned by the functions $\bar{x}_k$.

Now, this approximation also automatically entails a scheme for the generation of X_k triples from the input ρ_k. Since we have

$$\rho_i \;=\; \int R_i(\lambda)\, C(\lambda)\, d\lambda$$

$$\simeq \sum_{k=1}^{3} a_{ik} \int \bar{x}_k(\lambda)\, C(\lambda)\, d\lambda$$

$$= \sum_{k=1}^{3} a_{ik} X_k \tag{6}$$

then the input ρ_k are approximately given by a linear transformation of the X_k, and vice-versa:

$$\boldsymbol{X} \simeq QV^{-1}\boldsymbol{\rho}\,. \tag{7}$$

In section 5 we explore how well this approximation does in reproducing the X_k generated from an exact knowledge of the input color signal.

3 Metamerism and Color Signal Reconstruction

The problem with the above formulation is that while it does generate the "best" sensor *functions* from the actual R_k functions there is no guarantee that it generates the best tristimulus values X_k. In search of a better method, we consider instead an approximation of the original color signal by a truncated series and ask how well the tristimulus values are matched.

Suppose that $C(\lambda)$ is approximated by

$$C(\lambda) \simeq \sum_{i=1}^{3} c_i\, C_i(\lambda) \tag{8}$$

where $C_i(\lambda)$ are basis functions that do a good job of describing most color signals in nature. Then applying eqn. (1) we find

$$\rho_k \simeq \sum_{i=1}^{3} c_i \int C_i(\lambda)\, R_k(\lambda)\, d\lambda$$

$$\equiv \sum_{i=1}^{3} c_i b_{ik} \tag{9}$$

where for the ρ_k we use those values of the sensor responses developed from the exact $C(\lambda)$. Since we have chosen the functions $C_i(\lambda)$ and they are fixed, we know the matrix b_{ik} and it can be precalculated. Since the above system of equations is linear, we can easily solve for the set of weights c_i, which turn out to be just a linearly transformed version of the ρ_k.

In terms of minimizations, the approximation eqn. (8) amounts to a minimization

$$\text{Min} \int [\, C(\lambda) - \sum_{i=1}^{3} c_i\, C_i(\lambda)\,]^2\, d\lambda. \tag{10}$$

The linear transformation (9) can be thought of as a minimization

$$\text{Min} \sum_{k=1}^{3} (\rho_k - \sum_{i=1}^{3} c_i \, b_{ik})^2 \tag{11}$$

provided the matrix b_{ik} is nonsingular.

Now we can ask what values of X_k are produced from this approximation of the color signal. From eqn. (2) we know that the actual X_k are given approximately by

$$X_k \simeq \sum_{i=1}^{3} c_i \int C_i(\lambda) \, x_k \, d\lambda$$

$$\equiv \sum_{i=1}^{3} c_i c_{ik} \tag{12}$$

or, in matrix form,

$$\boldsymbol{X}^T \simeq \boldsymbol{c}^T E \tag{13}$$

Writing eqn. (9) as

$$\boldsymbol{c}^T = \boldsymbol{\rho}^T B^{-1} \tag{14}$$

and combining with the above, we have as our final approximation of the set of X_k the equation

$$\boldsymbol{X}^T \simeq \boldsymbol{\rho}^T B^{-1} E \tag{15}$$

Since matrices B and E are fixed by the choices of sensor filters and color-signal basis functions, we are left with a simple linear transformation from the ρ_k to the X_k that is fixed for any pixel, independent of the values of ρ_k. This is precisely the situation that obtains in respect to Method A, described in the last section. The difference lies in the definitions of B and E, which both incorporate the choice of the color signal basis $C_i(\lambda)$. We show below that with a good choice of basis set one can greatly improve on the values of X_k derived using Method A. Our approach is equivalent to the minimization (11), which determines the best basis function weights c_i given sensor responses ρ_k. Because we are not fitting an entire sensor response curve to the color-matching functions, the resulting, approximate X_k are better. As well, Method B, involving basis functions, generates the entire $C(\lambda)$ spectrum which can be examined for physical plausibility.

4 Color Signal Basis Sets

4.1 Fourier basis

As advocated in [20, 4], a set of three frequency-limited functions of wavelength can be used as a basis set. We can take as a simple set the first three Fourier functions, as follows:

$$F_1(\lambda) = 1$$

$$F_2(\lambda) = \sin[2\pi(\lambda - \lambda_{min})/(\lambda_{max} - \lambda_{min})]$$

$$F_3(\lambda) = \cos[2\pi(\lambda - \lambda_{min})/(\lambda_{max} - \lambda_{min})] \qquad (16).$$

Weighted sums of this set of functions can be expected to generate many physical chromaticities [4], but it is not clear whether such spectra correspond to natural color signals or are simply camera metamers that may be useful in generating RGB signals for use in graphics. In section 5 we test how well weighted sums of the above functions correspond to actual color signals.

4.2 Basis function products

A set of basis functions that has been used in connection with finite dimensional linear models of illumination and reflectance consists of product functions of a basis set for illumination times a basis set for reflectance [8, 2]. Suppose we use a basis set for illumination consisting of Judd *et al.*'s set of daylight bases derived from a principal component analysis of many illumination SPDs [12]. Denote by $E_i(\lambda)$ these illumination basis functions. Similarly, let $S_j(\lambda)$ be a basis set for reflectance. Brainard *et al.* [2] use the reflectance basis of Cohen [5], or alternatively basis functions they produced themselves using a large set of Munsell chips, again developed using a principal component analysis (*cf.* Maloney's analysis [16] of the large sample of natural reflectances obtained by Krinov [14]). Judd *et al.* modeled most daylights using just three to five illumination basis vectors; Cohen concluded that between three and six basis vectors were sufficient for modeling reflectance.

Here we are concerned with choosing just three product basis functions $P(\lambda)$ for modeling color signals; we select

$$P_1(\lambda) = E_1(\lambda)\, S_1(\lambda)$$

$$P_2(\lambda) = E_1(\lambda)\, S_2(\lambda)$$

$$P_3(\lambda) = E_2(\lambda)\, S_1(\lambda) \qquad (17)$$

and use as our illumination basis Judd *et al.*'s functions $E_i(\lambda)$. For the reflectance basis set, we follow Maloney [16] and carry out a principal component analysis on the Krinov catalogue of 370 natural reflectances [14]. [3] Since these reflectances are available in a limited range in the visible— 400nm through 650nm in steps of 10nm— we keep the analysis to only 26 samples over wavelength. In fact, since we use the Krinov reflectances to generate test color signals we use 26-component vectors throughout this work.

Judd *et al.*'s analysis of illumination data was a standard principal component analysis [11]. In this type of analysis one views the spectra as vectors in an N-dimensional vector space. Then as a first step one carries out a translation in the vector space by shifting each vector to a new origin, the mean vector of all the spectra. One then forms the variance-covariance matrix of all the vector components with each other over all the cases studied and diagonalizes this matrix. The vectors that result are ordered such that the first vector is in the direction that accounts for the maximum variability in the whole set of samples. The second vector is perpendicular to the first and accounts for the direction showing the next most important source of variability, and so on.

In modeling color signals as the product of a linear series in an illumination basis times a similar series in a reflectance basis, there can be uniqueness in the illumination and reflectance weights only up to an overall multiplicative factor because the color signal is formed by multiplication [8]. It then makes sense to keep to a weight of 1 for the first illumination basis function, since an overall choice makes no difference. Therefore using the mean vector plus a linear series for illumination, as provided by a standard principal component analysis, is just what is required. For reflectance, however, it makes more sense to derive basis vectors that are all orthonormal, since we are not setting the first weight to a special value and more importantly because for any new reflectance we do not have available a sample mean by which to translate the origin. Therefore, instead of a standard principal component analysis it seems reasonable to use a Karhunen-Loève analysis [22, p. 275], which does not translate by the mean and yields basis vectors that are all orthogonal. In its simplest form, the Karhunen-Loève transformation diagonalizes the raw component crossproduct matrix (the autocorrelation matrix) rather than the variance-covariance matrix (although the term can also refer to the standard principal component analysis). This amounts to viewing the origin, rather than the mean of the reflectance set, as a distinguished point. [4]

Using the product basis set P_i turns out to work quite well, better than the Fourier basis, for recovering color signals from ρ_k values, as shown in the section 5. This is due to the fact

[3] Actually we used only 342 of the Krinov spectra; we omitted those that were incomplete in the limited wavelength range utilized by Krinov and omitted three more (snows) that had reflectance values above 1.0.

[4] Alternatively, one can engineer the sample set such that the mean vector is zero [16].

that the P_i incorporate a good deal of statistical information on how natural color signals are formed.

4.3 Color signal space basis

A further basis set that suggests itself can be formed by carrying out a similar statistical analysis, but in the space of color signals itself. By forming a large number of synthetic color signals, as products of typical illumination spectra with natural reflectance spectra, we created a large data set to examine. We used the five standard Judd daylights [12] for correlated color temperatures between $4800^o K$ and $10000^o K$, multiplied by 342 reflectance spectra from the Krinov catalogue to obtain 1710 different synthetic, yet natural, color signals. We performed a Karhunen-Loève analysis on these spectra and derived a set of basis vectors, the first five of which are shown in Fig. 1. [5] Denote by $C_i(\lambda)$, $i = 1 \ldots 3$ the color signal space basis set to be used in eqn. (8). We show in the next section that these C_i perform the best of the three basis sets introduced here in mapping an input set of ρ_k back to a color signal.

It should be noted that any of the basis sets entails a possible gamut of XYZs and hence chromaticities that entirely covers the 1931 CIE chromaticity diagram. This is so because each set contains basis vectors with negative components and for each set some of the basis vectors produce negative values X_k. If the c_i are unrestricted, then any set of X_k can be generated. However, since we insist that any $C(\lambda)$ that we construct be nonnegative for every value of λ and since $\bar{x}_k(\lambda)$ is also nonnegative, the values of X_k are necessarily nonnegative also. Hence the chromaticities are restricted to being less than 1. In fact, because of the additive property of chromaticities, any positive $C(\lambda)$ also lies inside the spectrum locus (i.e. the chromaticities of all monochromatic signals).

The conditions $X_k \geq 0$ also provide a quick cue to whether a color signal is physically plausible. For from eqn. (13), that condition implies

$$c^T E \geq 0$$

so that possible values of c_i are restricted.

In the next section we look at results for Method A and the results using Method B for each of the three basis sets. For convenience, we refer to the methods as shown below:

[5] Buchsbaum [3] uses a Karhunen-Loève expansion as well, but diagonalizes in an RGB space, not with respect to the components of the color signal. See also [18].

Method A	method derived from Horn's analysis
Method B.F	basis function method using Fourier basis
Method B.P	using $E_i S_j$ product basis
Method B.C	using color signal space basis

5 Results

We use two different sensor response functions. First, we use the human cone responses given by Bowmaker and Dartnell [1], shown in Fig. 2. As well, for typical camera filters we used Kodak filters #25 (red), #58 (green) and #47B (blue), also shown in Fig. 2.

As a measure of accuracy of the color signal reconstruction, we use the CIELUV unit of distance (ΔE) in uniform color space [24]. Since a color signal is already an illumination multiplied by a reflectance, it is not necessary to multiply by the standard light D65 to obtain tristimulus values.

Here we consider one-by-one a series of color signals constructed synthetically. Since we do not have a complete image, it is not possible to adjust one color signal with respect to another to account for possible changes in luminance from pixel to pixel. Therefore, we normalize each color signal by setting the Y value to 100, as is done for individual illuminations in [24, p. 149]. The X_k given by eqn. (7) must also be normalized in the same way, i.e. by multiplication with $100/Y_{actual}$ where Y_{actual} is the unnormalized value of the luminance for the actual color signal. Similarly for the basis function methods, one should scale the tristimulus values calculated from eqn. (15) by Y_{actual}.

The set of synthetic color signals used to test each method consists of the five Judd standard daylights as well as the mean vector for daylight in Australia [6], each multiplied by each of the Krinov reflectance functions [14]. For the human cone response functions, the results are shown in Fig. 3 and Table 1.

	Mean	Median	StdDev
Method A	16.4	16.2	6.6
Method B.F	16.9	16.5	6.9
Method B.P	11.8	9.8	9.3
Method B.C	7.4	5.4	6.6

Table 1: Statistics using human cone response functions for 2052 typical color signals.

As can be seen, from poorest to best the methods are: Method B.F; Method A; Method B.P; Method B.C. The best mean uniform color space distance is 7.4 units. For the color

signal space basis (Method B.C), 2% of the approximate color signals were not included in the histogram analysis (Fig. 3) because at least one vector component turned out to be negative. Of the 2052 sample color signals examined, 37 had at least one negative component. Of these, the average number of negative components was 2, the median was 2 and the maximum was 4. Another way of dealing with reconstructed signals with negative components, rather than simply omitting them, would be by adding metameric black signals [24, p. 187], but we do not address this option here.

For each of the other basis function methods, the histograms shown also omit any signals with negative components. Signals that must be omitted are strongly correlated amongst the three basis function methods, Methods B. For Method A, it is not possible to screen unphysical color signals in this way. The fact that the cone functions are close to being just a linear transformation of the color-matching functions means that the non-basis method, Method A, works not too badly.

The relative ranking of the basis function methods persists when we use more realistic sensor functions. However, in this case Method A works not at all well (up by an order of magnitude in CIELUV units), presumably because the sensor functions are far from being a linear combination of eye functions. Using the camera filters, we arrive at the results in Fig. 4 and Table 2, where only the basis function methods are shown. From poorest to best the methods are once again Method B.F; Method B.P; Method B.C, with a best average CIELUV distance of 5.9 units. For all the methods the standard deviations are seen to be quite wide.

	Mean	Median	StdDev
Method B.F	22.8	20.5	14.3
Method B.P	6.6	5.2	5.4
Method B.C	5.9	4.5	5.0

Table 2: Statistics using camera response functions for 2052 typical color signals.

To get an idea of what these figures mean, consider applying the color-signal-space basis model to a color signal composed of the Australian illumination spectrum multiplied by the first principal component vector in Cohen's analysis of reflectances. We chose these spectra because neither had anything to do with the development of the color-signal-space basis set. Applying eqn. (15) and using the camera sensors, we arrive at a CIELUV error of 3.6 units. Fig. 5 shows the original color signal and the approximation derived by the algorithm. Their agreement is striking.

For this 'typical' *natural* color signal, we can repeat the color display test performed on the artificial reflectances in the Introduction. For the actual color signal the tristimulus

values are $X_k = (98.55, 100.00, 95.59)$ so that the chromaticity is $(x, y) = (0.3350, 0.3400)$. This color is very close to an ideal white. The camera signal is $\rho_k = (1.084, 1.013, 1.006)$ which displays as a pale pink—the camera does not see what the eye does. However, Method B.C maps these ρ_k back to tristimulus values $X_k = (97.89, 100.35, 92.90)$, with chromaticity $(x, y) = (0.3362, 0.3447)$, i.e., approximately back once again to the white the eye sees. The resulting CIELUV error is $\Delta E = 3.59$.

The average error using the color-signal-space basis and camera sensors was $\Delta E = 5.9$ units. For comparison, Stamm [19] showed that the average allowable tolerance in a typical printing industry application was $\Delta E = 6$ units, with standard deviations ~3-4 (in the related CIELAB system). We are considering just the first step in a process that might accumulate further errors; however, Stamm studied a printing duplication application, not the more difficult problem of accurately reproducing actual color. For this more difficult problem, the results presented here seem adequate (*cf.* the comparable and larger numbers in [18, 21]).

6 Conclusions

While the method derived from Horn's work (Method A) performs reasonably well when applied to sensor RGBs that result from sensor sensitivities close to human responses, we have shown that that method is unreliable when the response curves are far from being linear transforms of human cone responses.

In contrast, our method based on basis functions works well using either cone response functions or camera response functions. The best model results from using a set of basis functions derived from a principal component analysis in color signal space, since these basis functions recover the most "natural" color signal that is a camera metamer to the input RGB.

Since the basis-function method boils down to a simple linear transformation that is independent of the actual values of ρ_k at any pixel, the method achieves exact reproduction of XYZ when the camera response curves used are exactly a linear transform of the color-matching functions, just as the Method A does. However, the basis-function method is superior when the linear-transform condition does not hold.

Method B also produces the entire color signal spectrum as an intermediate step. We found the color-signal basis derived from the statistical properties of natural color signals to be the one most likely to avoid color-signal spectra containing negative components. The approximate color signal SPD can be examined for unphysical components, and is of interest in its own right for other applications. For example, we note that having the color-

signal spectrum available would allow the direct application of the Dichromatic Reflection Model [13] for the separation of surface- and body-reflection components. As we shall explore elsewhere, carrying out the separation in the c_i space should allow one to recover the spectrum of the incident illumination.

7 Acknowledgements

M.S. Drew is indebted to the Centre for Systems Science at Simon Fraser University for partial support; B.V. Funt thanks both the CSS and the Natural Sciences and Engineering Research Council of Canada for their support.

8 Figure captions

Figure 1. The first five color signal space basis vectors derived from a Karhunen-Loève analysis of 1710 synthetic natural color signals.

Figure 2. Sensor response functions for human cones and typical camera sensors. Red: solid lines; Green: small dashes; Blue: longer dashes. In each case, the camera curves peak to the right (longer wavelengths side) of the cone curves.

Figure 3. Histograms showing frequency over ΔE for recovered color signal compared to actual color signal for each method, using human cone sensor response functions. The number of synthetic spectra tested was 2052. For the basis function methods the percentages do not add up to 100% because cases were omitted if one or more components of the recovered color signal were negative. Percentages retained were: Fourier: 88.7%, Product: 90.4%, Color signal space basis: 98.2%.

Figure 4. Frequencies using camera sensor response functions for 2052 synthetic spectra. The method derived from Horn had an average ΔE of 133 and is omitted. For the other methods the percentages of cases retained were: Fourier: 91.5%, Product: 97.7%, Color signal space basis: 98.1%.

Figure 5. Sample test color signal: mean vector for Australian daylight multiplied by first principal component vector for Munsell chip reflectances, as given by Cohen.

References

[1] J.K. Bowmaker and H.J.A.Dartnell. Visual pigments of rods and cones in the human retina. *J. Physiol.*, 298:501–511, 1980.

[2] D.A. Brainard, B.A. Wandell, and W.B. Cowan. Black light: how sensors filter spectral variation of the illuminant. *IEEE Trans. Biomed. Eng.*, 36:140–149 and 572, 1989.

[3] G. Buchsbaum and A. Gottschalk. Trichromacy, opponent colours coding and optimum colour information transmission in the retina. *Proc. R. Soc. Lond. B.*, 220:89–113, 1983.

[4] G. Buchsbaum and A. Gottschalk. Chromaticity coordinates of frequency-limited functions. *J. Opt. Soc. Am. A*, 1:885–887, 1984.

[5] J. Cohen. Dependency of the spectral reflectance curves of the munsell color chips. *Psychon. Sci.*, 1:369–370, 1964.

[6] E.R. Dixon. Spectral distribution of australian daylight. *J. Opt. Soc. Am*, 68:437-450, 1978.

[7] B.V. Funt and J. Ho. Color from black and white. In *Proceedings of the Second International Conference on Computer Vision, Tarpon Springs Dec 5-8*, pages 2–8. IEEE, 1988. and *Int. J. Computer Vision*, 3:109-117, 1989.

[8] J. Ho, B.V. Funt, and M.S. Drew. Separating a color signal into illumination and surface reflectance components: Theory and applications. *IEEE Trans. Patt. Anal. and Mach. Intell.*, 1990. In press.

[9] B.K.P. Horn. Exact reproduction of colored images. *Comp. Vision, Graphics and Image Proc.*, 26:135–167, 1984.

[10] R.W.G. Hunt. *The Reproduction of Color.* Wiley, 3rd edition, 1975.

[11] I.T. Jolliffe. *Principal Component Analysis.* Springer-Verlag, 1986.

[12] D.B. Judd, D.L. MacAdam, and G. Wyszecki. Spectral distribution of typical daylight as a function of correlated color temperature. *J. Opt. Soc. Am.*, 54:1031–1040, August 1964.

[13] G.J. Klinker, S.S. Shafer, and T. Kanade. The measurement of highlights on color images. *Int J. Comp. Vis.*, pages 7–32, 1988.

[14] E.L. Krinov. Spectral reflectance properties of natural formations. *Technical Translation TT-439, National Research Council of Canada*, 1947.

[15] R.L. Lee Jr. Colorimetric calibration of a video digitizing system: algorithm and applications. *Color Research and Applications*, pages 180–186, 1988.

[16] L.T. Maloney. Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *J. Opt. Soc. Am. A*, 3:1673–1683. 1986.

[17] L.T. Maloney and B.A. Wandell. Color constancy: a method for recovering surface spectral reflectance. *J. Opt. Soc. Am. A*, 3:29–33, 1986.

[18] G.W. Meyer. Wavelength selection for synthetic image generation. *Comp. Vision, Graphics, and Image Proc.*, 41:57–79, 1988.

[19] S. Stamm. An investigation of color tolerance. In *Technical Association of the Graphic Arts Conference*, pages 157–173, 1981.

[20] W.S. Stiles, G. Wyszecki, and N. Ohta. Counting metameric object-color stimuli using frequency-limited spectral reflectance function. *J. Opt. Soc. Am.*, 67:779–784, 1977.

[21] M.C. Stone, W.B. Cowan, and J.C. Beatty. Color gamut mapping and the printing of digital color images. *ACM Trans. Graphics*, pages 249–292, 1988.

[22] J.T. Tou and R.C. Gonzalez. *Pattern Recognition Principles.* Addison-Wesley, 1974.

[23] B.A. Wandell. The synthesis and analysis of color images. *IEEE Trans. Patt. Anal. and Mach. Intell.*, PAMI-9:2–13, 1987.

[24] G. Wyszecki and W.S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulas.* Wiley, New York, 2nd edition, 1982.

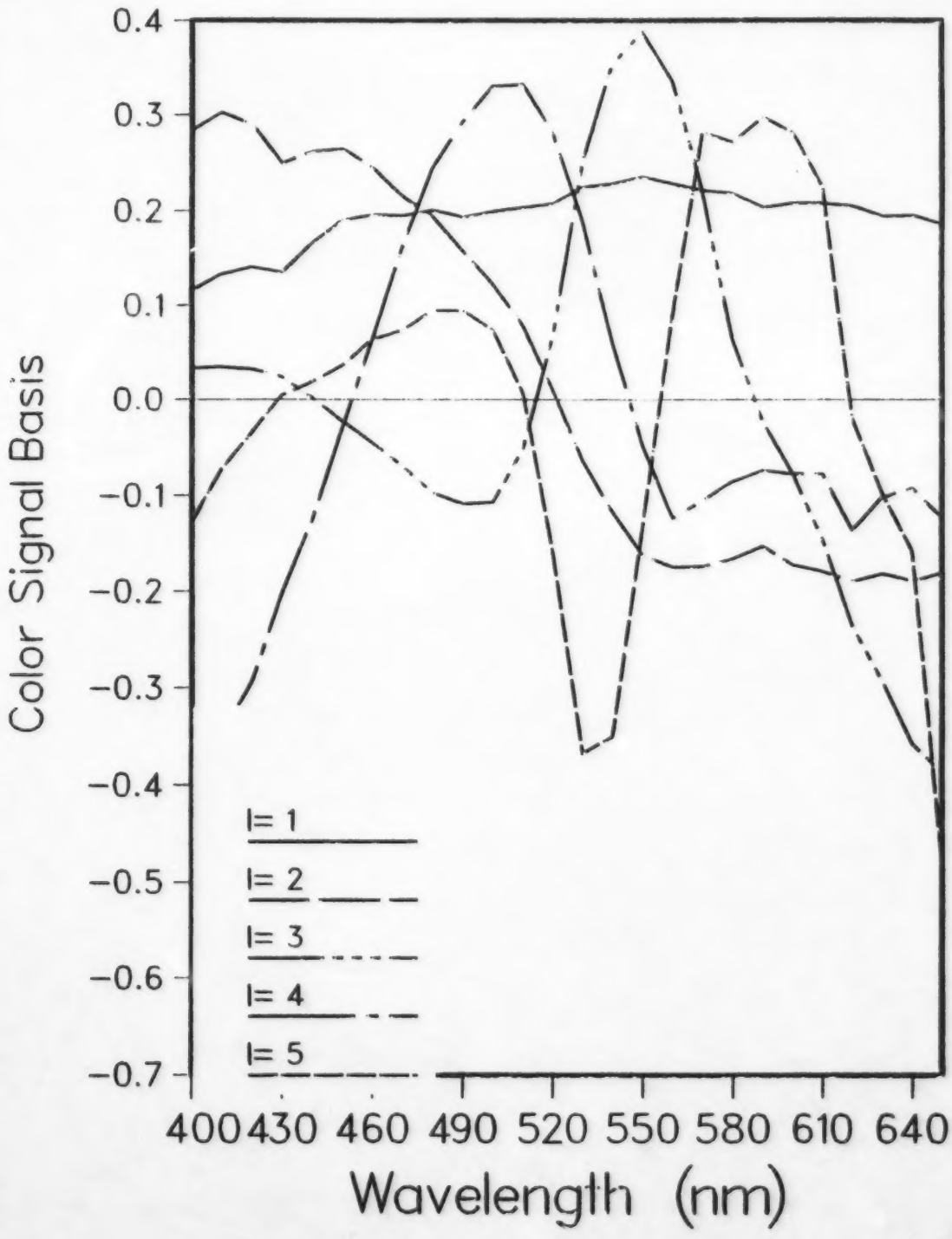
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
−0.7
Color Signal Basis
I= 1
I= 2
I= 3
I= 4
I= 5
400 430 460 490 520 550 580 610 640
Wavelength (nm)

Fig. 1

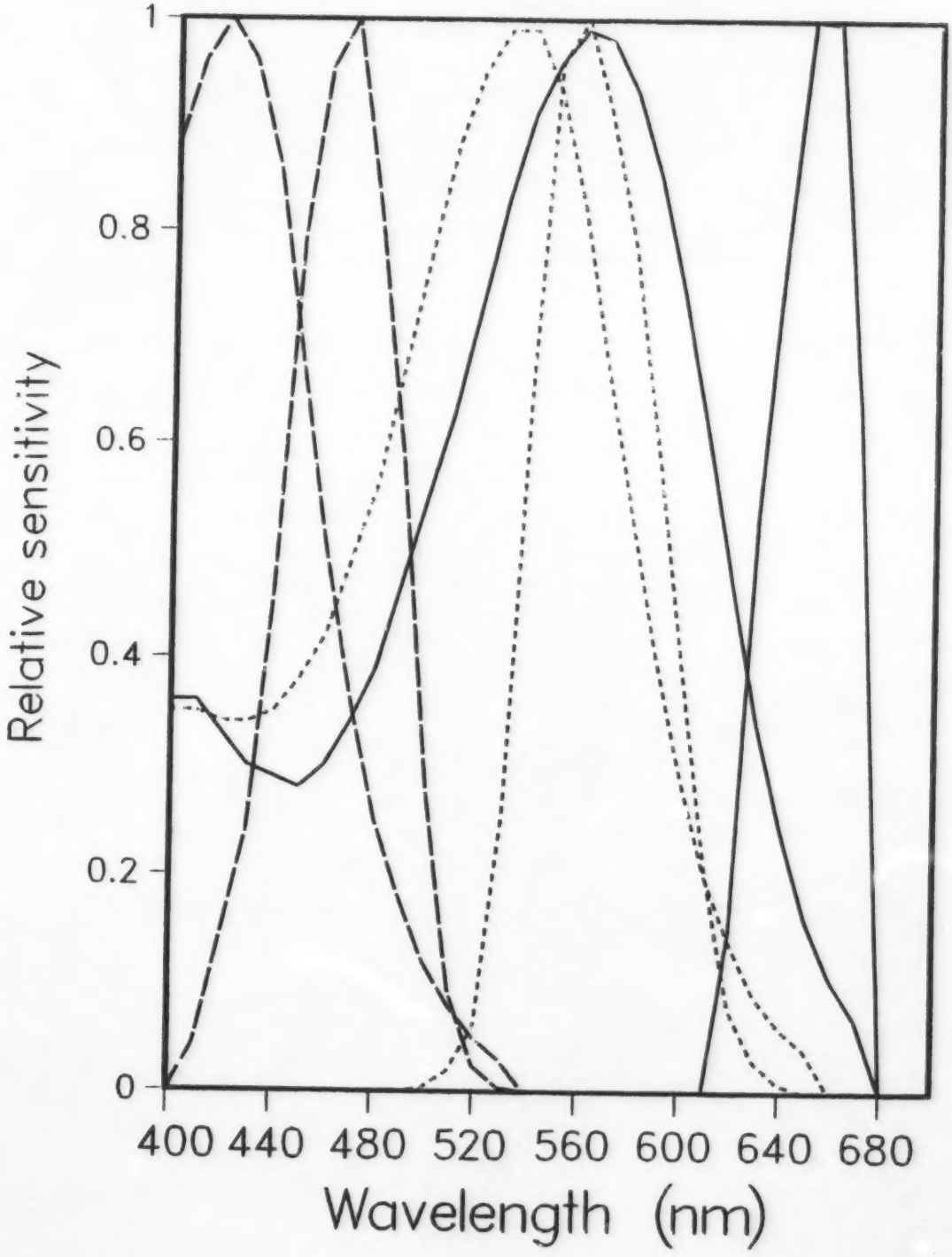
1
0.8
0.6
0.4
0.2
0
Relative sensitivity
400 440 480 520 560 600 640 680
Wavelength (nm)

Fig. 2

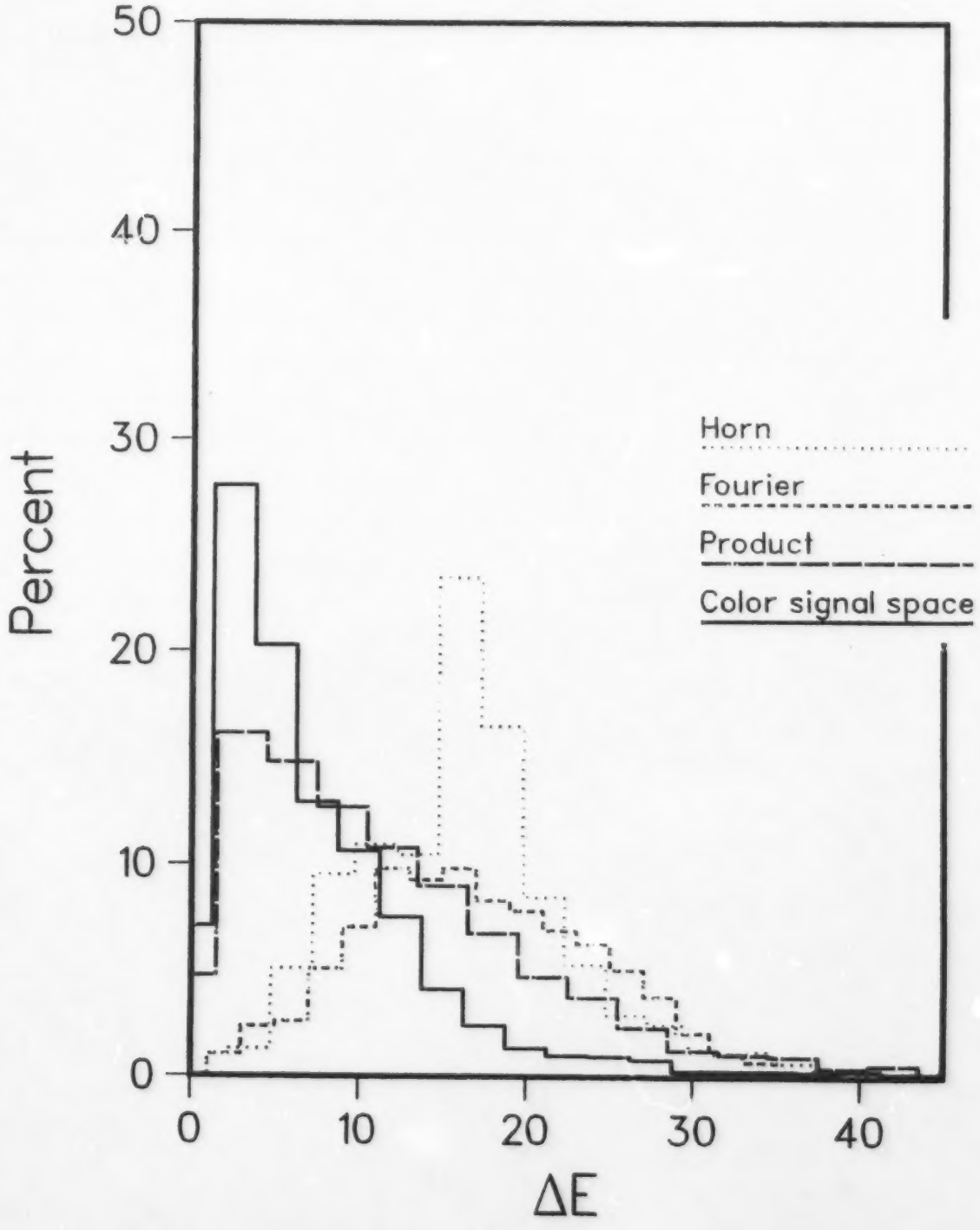
50
40
30
20
10
0
Percent
0
10
20
30
40
ΔE
Horn
Fourier
Product
Color signal space

Fig. 3

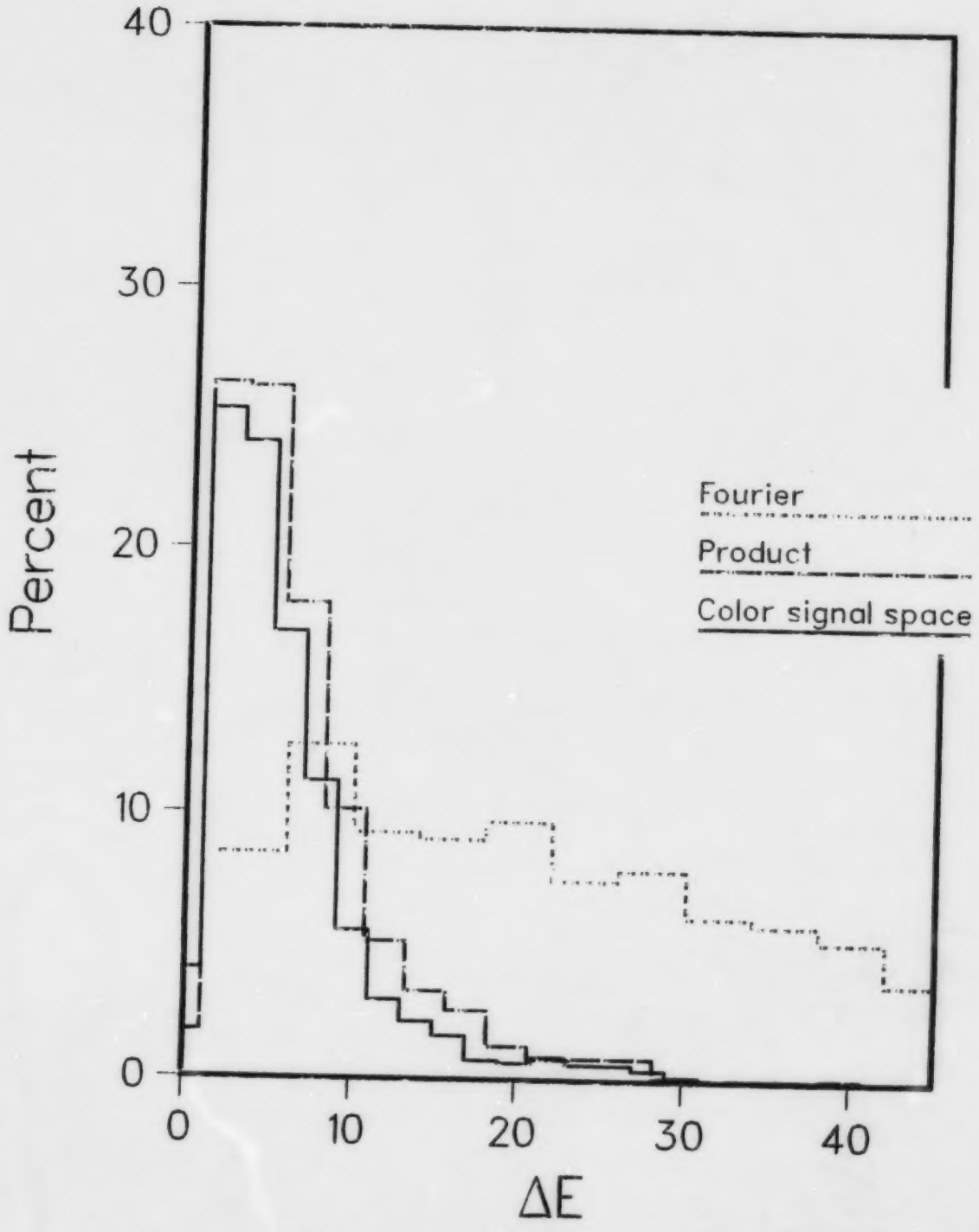
40
30
20
10
0
Percent
0
10
20
30
40
ΔE
Fourier
Product
Color signal space

Fig. 4

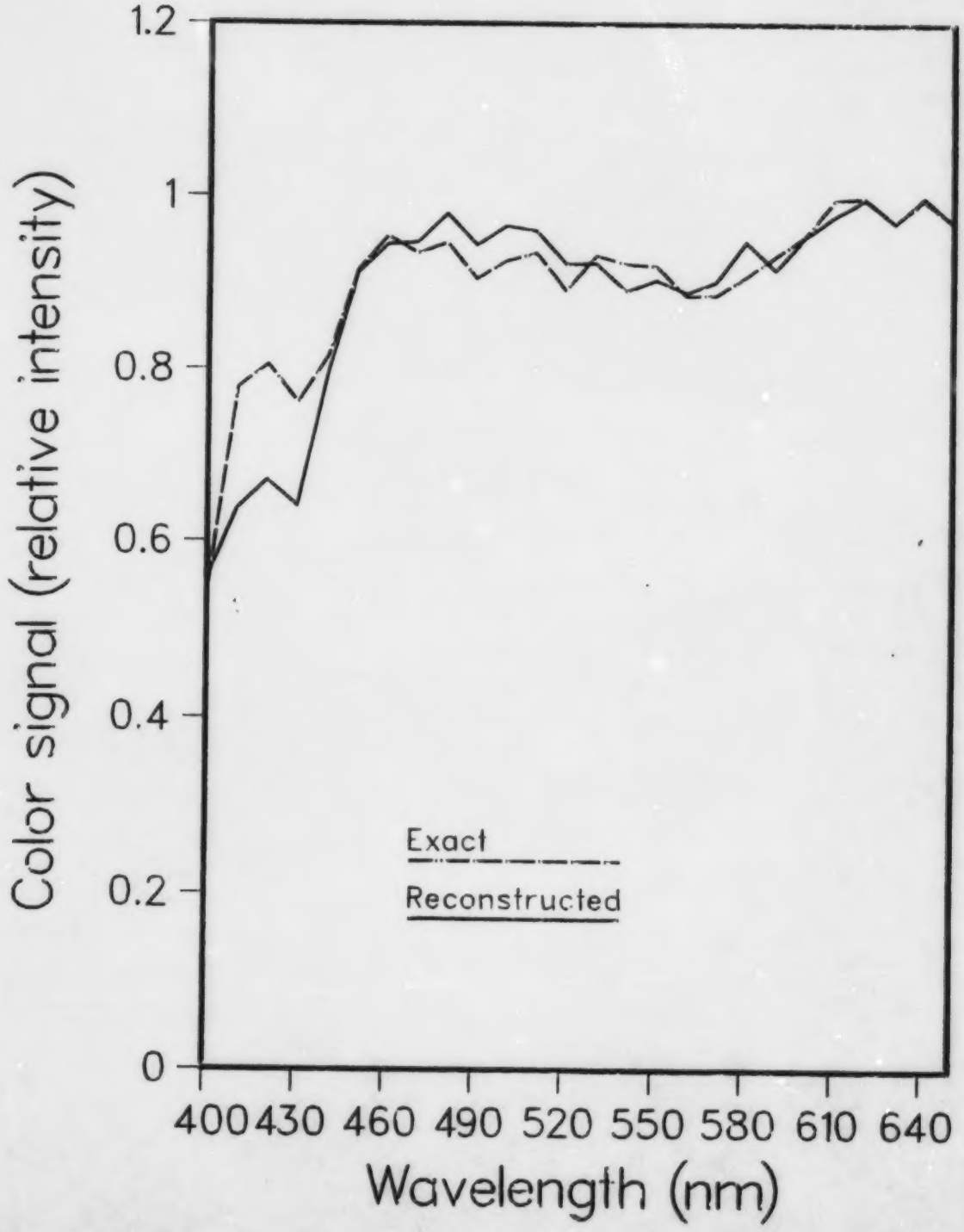
1.2
1
0.8
0.6
0.4
0.2
0
Color signal (relative intensity)
Exact
Reconstructed
400 430 460 490 520 550 580 610 640
Wavelength (nm)

Fig. 5

Centre for Systems Science

• Simon Fraser University established the **Centre for Systems Science** as a provincial centre for advanced study. The objective of CSS is twofold: coordinate and combine advanced systems research at SFU, and encourage application of research results within industry in order to further the technological development of British Columbia.

• CSS has assembled three research groups to streamline its efforts. The orientation of these groups is as follows: intelligent systems, microelectronics, and computer and communications systems. Because of shared research interests, this latter group merges with LCCR (**Laboratory for Computer and Communications Research**) for the publication of technical reports.

• CSS's three groups involve over 80 skilled faculty researchers principally from SFU's Schools of Engineering Science and Computing Science, but also draw substantially on Mathematics, Physics, Geography, Philosophy, Linguistics, Psychology, Communication, and Kinesiology.